

# Detección de edificios en imágenes urbanas

Federico Cristina, Sebastián H. Dapoto, Verónica Artola, Claudia Russo, Graciela Viegas

Instituto de Investigación en Informática LIDI (III-LIDI), Facultad de Informática,  
Universidad Nacional de La Plata, 50 y 120 2do piso, La Plata, Argentina  
Unidad de Investigación N° 2 del Instituto de Estudios del Hábitat (IDEHAB), Facultad de  
Arquitectura y Urbanismo, Universidad Nacional de La Plata  
47 N° 162, La Plata, Argentina.

{fcristina, sdapoto, vartola, crusso} @lidi.info.unlp.edu.ar  
{gachiviegas} @yahoo.com.ar

**Abstract.** El estudio de sectores reducidos representativos de un área urbana homogénea - mosaicos urbanos - permite estimar la necesidad de aplicar sistemas energéticos alternativos e implementar estrategias de uso eficiente de la energía. Para dicho estudio es necesario llevar a cabo el relevamiento de mosaicos urbanos a fin de identificar las edificaciones. Realizada de forma manual, esta tarea es tediosa y consume gran cantidad de tiempo. Este trabajo presenta el análisis de las técnicas de detección de edificios de forma automática/semiautomática sobre imágenes aéreas y satelitales de libre acceso, y sus resultados preliminares.

**Keywords:** detección automática, mosaicos urbanos, segmentación, imágenes aéreas

## 1. Introducción

El presente trabajo es parte de un proyecto de investigación cuyo objetivo general busca elaborar una metodología de análisis, aplicada a sectores urbanos reducidos, que permita visualizar y realizar estimaciones orientadas a la sustitución de fuentes de energía tradicionales por renovables y a la implementación de estrategias de uso eficiente de la energía.

La unidad de análisis está constituida por sectores urbanos reducidos, definidos como “mosaicos urbanos”. Los mismos son representativos de áreas homogéneas del Partido de La Plata respecto a la ocupación y uso del suelo, la cobertura de redes de servicios, el uso predominante del suelo (tal como residencial o terciario), y las características tipológicas edilicias. El estudio de mosaicos urbanos permite obtener indicadores que posibilitan el diagnóstico de la situación existente para la aplicación de fuentes sustitutas y medidas correctivas de uso eficiente de la energía, minimizando los impactos negativos detectados.

Para el relevamiento de los mosaicos se parte del análisis de imágenes aéreas o satelitales. En un procedimiento de análisis manual, un operador calificado realiza el relevamiento por observación y sintetiza la información manualmente en programas de diseño asistido por computadora (CAD), planillas de cálculo y/o sistemas de información geográfica (GIS).

La práctica del relevamiento manual presenta como mayor dificultad el tiempo de trabajo requerido y la necesidad de realizar verificaciones en el sector analizado.

Con el apoyo de técnicas ad-hoc de procesamiento de imágenes, es posible facilitar la tarea de relevamiento automatizando los procesos de detección de edificaciones.

El presente trabajo tiene como objetivo explorar las técnicas de segmentación de imágenes existentes para la detección de edificaciones en imágenes aéreas de acceso libre buscando la manera de facilitar el relevamiento de áreas urbanas representativas.

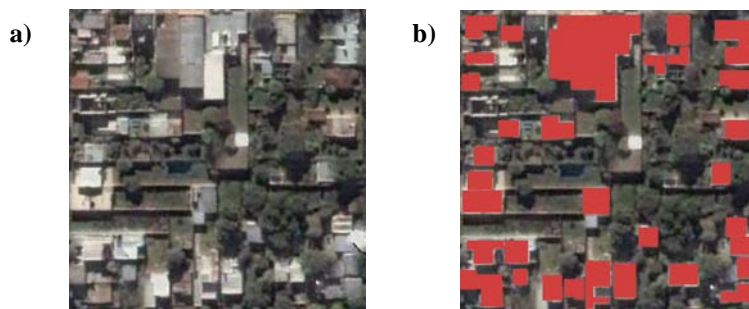
A fin de contrastar los resultados obtenidos mediante las técnicas analizadas, se toman como referencia los resultados del relevamiento de mosaicos urbanos realizados en forma manual en investigaciones previas [1] (Figuras 3.1 y 3.2).

## 2. Técnicas de segmentación

Las imágenes urbanas son complejas y presentan un gran número de objetos. Por lo general no tienen un fondo definido, por lo que la detección de objetos en ellas se torna dificultosa inclusive para el ojo humano.

En la actualidad existe una gran variedad de imágenes de este tipo, las cuales varían en calidad, resolución y en el modo en que fueron obtenidas. Por ello, hay una importante cantidad de técnicas a utilizar en esta área, ajustándose cada una de éstas a las distintas necesidades y tipos de imágenes. En el mercado es posible encontrar servicios que ofrecen imágenes multiespectrales y modelos de superficie que ayudan a realizar un análisis de las mismas. Sin embargo, estos servicios presentan un costo prohibitivo para el usuario común.

El presente trabajo se encuentra centrado en imágenes de acceso libre, en particular, en las obtenidas de la aplicación Google Earth<sup>1</sup>. Se exponen además pruebas adicionales sobre imágenes aéreas monocromáticas obtenidas años atrás. Debido a este desfase en el tiempo, existe una leve diferencia en los objetos a detectar entre las muestras.

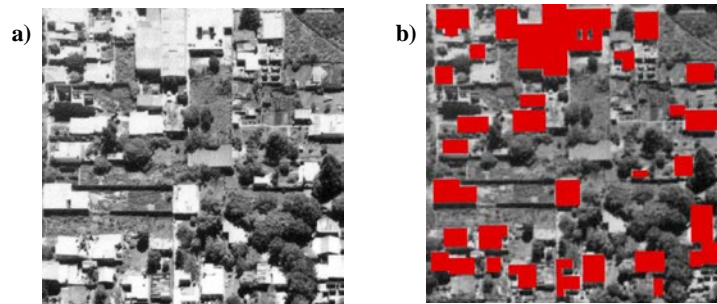


**Fig. 1. a)** Imagen satelital. Barrio La Loma, La Plata **b)** Relevamiento manual

Muchos trabajos se han realizado sobre distintos métodos para solucionar el problema de detección de edificios en imágenes urbanas. Algunos abordan el problema de la detección utilizando técnicas de detección de bordes en imágenes satelitales [2, 3, 12], mientras que otros usan gradiente morfológico [5]. Existe también una gran cantidad de trabajos que se encargan de imágenes satelitales multiespectrales [4], o con imágenes aéreas de alta resolución [10]. Por último, ciertos trabajos se basan en imágenes pancromáticas [11].

---

<sup>1</sup> Google Earth 5.0 Freeware - Google Inc. - [earth.google.com](http://earth.google.com)



**Fig. 2.** a) Imagen de aérea. Barrio La Loma, La Plata b) Relevamiento manual

La mayor dificultad que presentan las imágenes utilizadas en el presente trabajo es la baja calidad de las mismas. Como problemas principales, es posible enumerar: oclusión, ruido, errores de la imagen, falta de ortorectificación, sombras y confusión de suelo con edificios y viceversa.

Las técnicas basadas en detección de bordes no dieron buenos resultados aplicadas a las imágenes sobre las que se trabajó. Por otro lado, las pruebas con gradiente morfológico y Canny arrojaron resultados confusos debido a la cantidad de objetos presentes en la imagen y a los cambios de tonos presentes en un mismo objeto.

Por esta razón se orientó la investigación a la utilización de técnicas que trabajan con las tonalidades de gris y no con los bordes de los objetos. Las primeras pruebas realizadas fueron de segmentación por textura.

Este trabajo se enfocó específicamente en la detección de bloques con edificaciones. Estos pueden pertenecer a más de un edificio, pero dicha información, por el momento, no se ha tenido en cuenta.

### 3. Técnicas aplicadas

#### 3.1 Segmentación por textura

No existe una definición universal de textura. En este trabajo se asume que una textura es esencialmente una interacción estadística de píxeles, es por esto que se aborda el problema de analizarlas y clasificarlas por medios estadísticos.

Se asume que una imagen posee una resolución que puede contener varias texturas dependiendo de dicha resolución. En el follaje de los árboles a gran escala la textura predominante seguramente sea la flora verde contrastada con el suelo, la sombra y los edificios; mientras que en una escala mucho menor la textura predominante podría ser las hojas de un árbol.

En la etapa de caracterización se extrae un vector de características descriptivas de cada textura, que la identifica unívocamente. Cada clase no tiene un vector de características único, sino que tiene una familia de vectores que son muy similares entre sí pero que no son iguales. Estos tienen características como media, varianza y otras medidas de mayor orden muy cercanas entre sí. Se busca encontrar un vector que esté lo suficientemente diferenciado de los vectores de las otras clases.

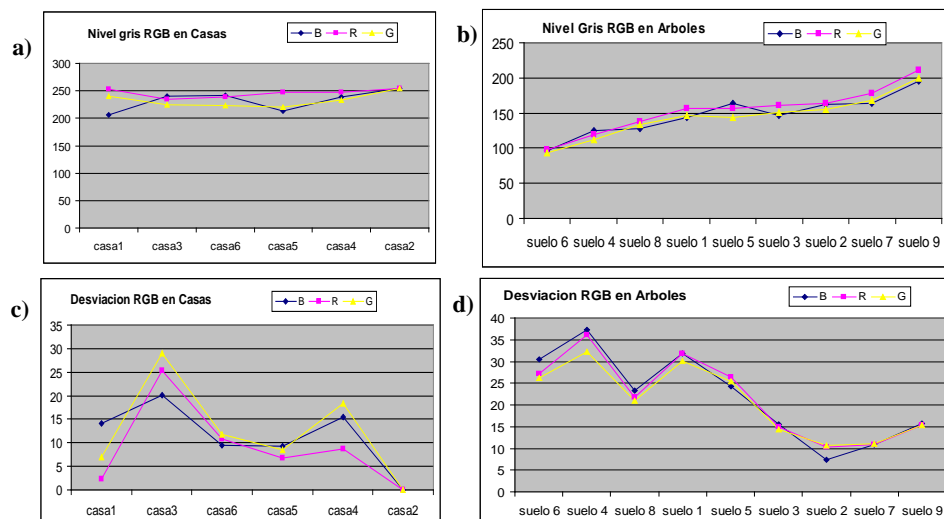
La etapa de caracterización es tal vez una de las más difíciles de llevar a cabo. Extraer vectores de buena calidad, es decir, que contengan información relevante sobre la entidad, implica que la posterior clasificación sea más sencilla de realizar. Esto se debe a que no se dará lugar a ambigüedades entre vectores ya que ningún

vector se solapará con otro. De otra manera se mezclarían, implicando eventuales clasificaciones erróneas. Lamentablemente en la práctica es muy difícil conseguir familias de vectores que cumplan con esta característica.

Se tomaron muestras de los objetos casas y árboles y se obtuvieron valores de diferentes descriptores en dichas muestras.

La elección de los descriptores más significativos debe ser rigurosa. No deberían incluirse aquellos que no aportan información realmente útil, ya que en las etapas posteriores, la redundancia agregaría complejidad al clasificador pero no mejoraría los resultados.

Luego de analizar la caracterización se consideró que los descriptores más representativos fueron: desviación y media. En las figuras 4 se exhiben ejes cartesianos, los cuales agrupan valores de dichos momentos para cada canal RGB

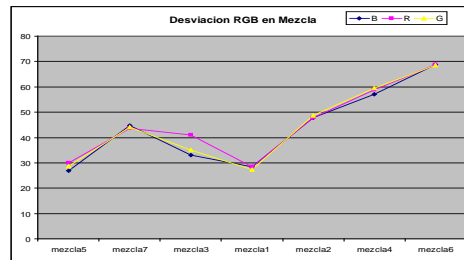


**Fig. 3. a) Media de gris en casas b) Media de gris en árboles c) Desviación en casas y d) Desviación en árboles**

La siguiente etapa implica la utilización de los descriptores antes escogidos. Para ello se realizó un barrido de la imagen con una ventana sobre la cual se calcularon los momentos escogidos como representativos. Esto significa que por cada vez que la ventana se deslizó de una posición a la siguiente se calculó, en la región abarcada, la media de nivel de gris y la desviación.

El tamaño de la ventana fue 20x20 y fue escogido después de realizar una serie de pruebas que arrojaron como óptimo el tamaño previamente mencionado.

Durante el barrido, la ventana puede abarcar uno o varios objetos. Este último caso requiere tomar nuevas muestras y calcular los valores de los descriptores como se efectuó con los árboles y las casas. De los resultados arrojados, el más representativo fue el valor de la desviación, ya que fue muy elevado con respecto al valor de los árboles y casas. Estos resultados son expuestos en la figura 3.5. Se puede observar que existe un límite inferior, en los tres canales alrededor del valor 25, y este valor es el utilizado como umbral para determinar si una región posee un solo objeto o varios.

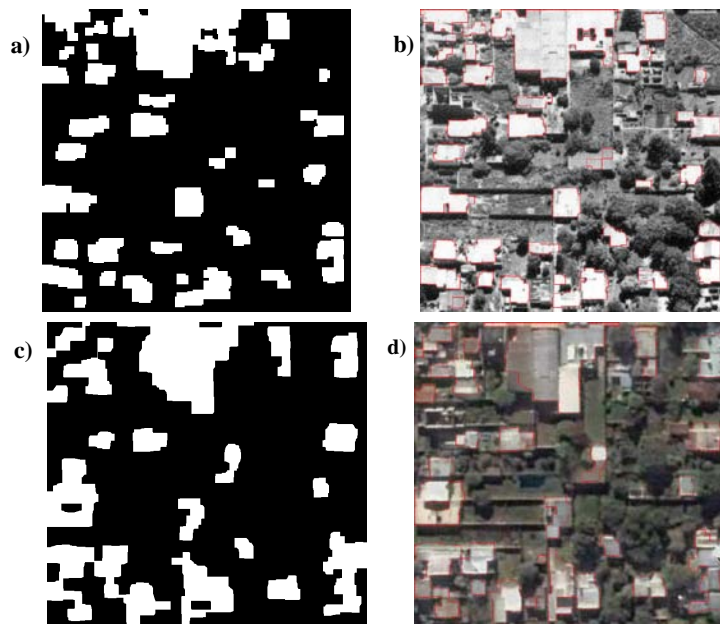


**Fig. 4.** Desviación en regiones con más de un objeto

A medida que se avanza en el barrido se chequea si la región en donde se encuentra la ventana se puede clasificar. Si los píxeles pertenecen a un solo objeto esto será posible, en caso contrario será necesario dividir la ventana y volver a chequear en las subimágenes. Este proceso se avanza recursivamente hasta lograr una clasificación.

Cuando en la ventana se encuentra un solo objeto, se analizan los descriptores de media y desviación nuevamente (figuras 3.3 y 3.4). Para clasificar una determinada región y determinar a qué grupo pertenece se utiliza umbralización, ya que la misma es una técnica simple y rápida.

Luego de realizar estas pruebas y analizar los resultados se propuso simplificar el proceso y directamente umbralizar la imagen por tonalidad de gris. De acuerdo a los resultados obtenidos se determinó que es posible descartar por completo el procedimiento de detección por texturas.



**Fig. 5** a) y b) Máscara y segmentación superpuesta con texturas en imágenes aéreas c) y d) Máscara y segmentación superpuesta con texturas en imágenes satelitales

### 3.2 Segmentación por umbralización

La umbralización es una técnica de segmentación empleada cuando hay una clara diferencia entre los objetos a extraer respecto del fondo de la escena. Los principios

que rigen son la similitud entre los píxeles pertenecientes a un objeto y sus diferencias respecto al resto. Por tanto, la escena debe caracterizarse por un fondo uniforme y por objetos parecidos. La dificultad radica en encontrar el valor correcto del umbral  $T$ .

### 3.2.1 Umbralización Otsu

Si bien hay diferentes métodos para hallar un umbral, la mayoría de ellos no proporcionan buenos resultados cuando se trabaja con imágenes del mundo real debido a la presencia de ruido, histogramas planos o una iluminación no adecuada. Por el contrario, el método de Otsu fue uno de los mejores métodos de selección de umbral para imágenes reales.

El método de Otsu, selecciona el umbral óptimo maximizando la varianza entre clases mediante una búsqueda exhaustiva. A medida que el número de clases de una imagen aumenta, el método de Otsu necesita mucho más tiempo para seleccionar un umbral multinivel adecuado.

La importancia del método de Otsu radica en que es automático, es decir, no necesita supervisión humana ni información previa de la imagen antes de su procesamiento.

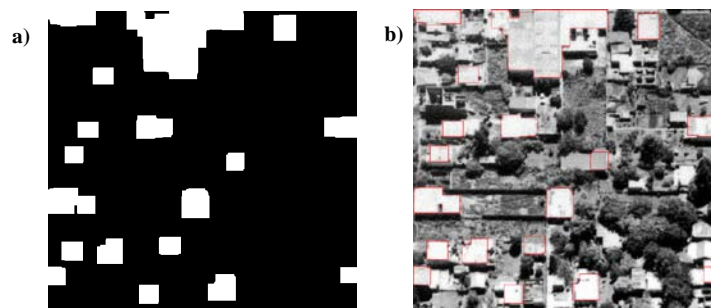
Para la finalidad de este trabajo se considera como fondo al suelo presente en la fotografía. Naturalmente el suelo puede poseer distintas texturas, tierra, pasto o tener algún tipo de material como baldosas, cemento, etc.

Los objetos presentes en la imagen pueden ser árboles, construcciones o sombras. En esta primera etapa se detectan únicamente las construcciones, por esto pertenecerán al mismo grupo todos aquellos que no sean edificios.

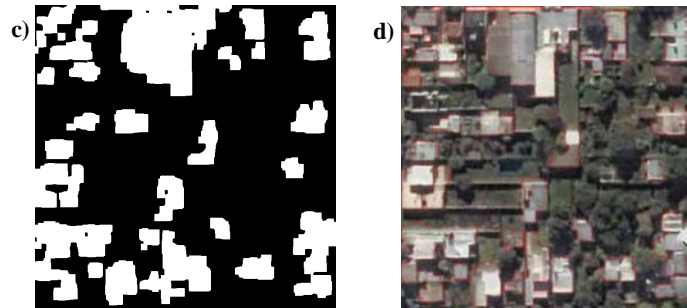
Observando las imágenes 3.1.a) y 3.2.a) se puede notar a simple vista que el suelo con pasto se asemeja a los árboles; y el suelo con baldosas u otro material se asemeja a las construcciones. Además, un dato útil es que las sombras tienen una tonalidad muy baja de gris que no se encuentra en los demás objetos.

Si bien las edificaciones poseen, en su mayoría, una alta tonalidad de gris, existen construcciones con techos más oscuros que se confunden con el suelo. En la mayoría de los casos tratados estas construcciones son pocas en relación a la cantidad de edificios.

El resultado de la umbralización de la imagen 3.1 se observa en la figura 3.8; y el de la 3.2 en la figura 3.9, en las cuales se puede observar la debilidad de esta técnica: las construcciones oscuras no fueron detectadas.



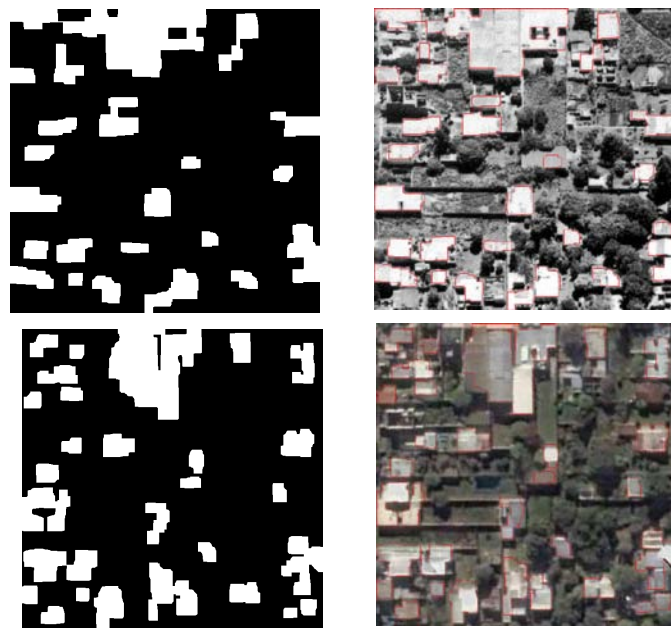




**Fig. 6** a) y b) Máscara y segmentación superpuesta con Otsu en imágenes aéreas c) y d) Máscara y segmentación superpuesta con Otsu en imágenes satelitales

### 3.2.2 Umbral a mano

El algoritmo de Otsu calcula un umbral óptimo de forma automática. Otra posibilidad es darle al usuario la capacidad de elegir el umbral a su conveniencia. Así se apuntaría a la creación sistemas semiautomáticos donde el usuario interactúa con el sistema formando parte del proceso de segmentación. De esta forma se pueden detectar construcciones que con Otsu no era posible. Por otro lado se puede observar que siguen existiendo algunas detecciones erróneas.

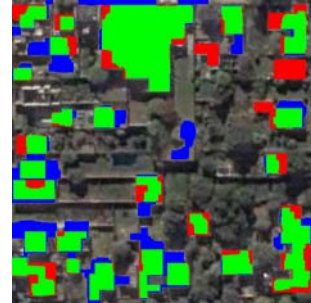


**Fig. 7** a) y b) Máscara y segmentación superpuesta con umbralización manual en imágenes aéreas c) y d) Máscara y segmentación superpuesta con umbralización manual en imágenes satelitales

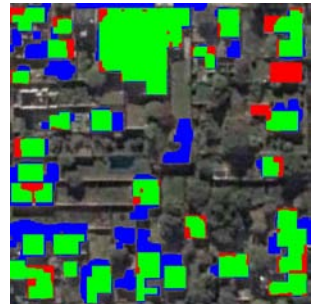
## 4. Estimación de error

### 4.1.1 Imágenes satelitales

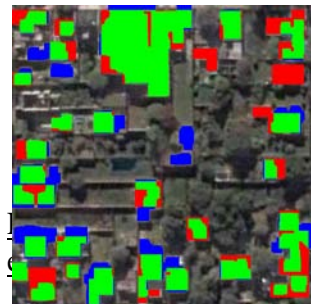
Texturas	Detectado	No Detectado	% Error
Construido	329831	33908	12.2%
No Construido	25305	93274	



Umbral Otsu	Detectado	No Detectado	% Error
Construido	315160	48579	13.4%
No Construido	16187	102392	

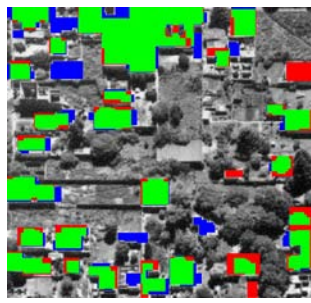


Umbral manual	Detectado	No Detectado	% Error
Construido	337085	26654	11.7%
No Construido	30011	88568	

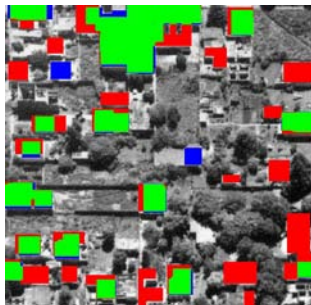


### 4.1.2 Imágenes aéreas

Texturas	Detectado	No Detectado	% Error
Construido	240785	15192	9%
No Construido	14281	56348	

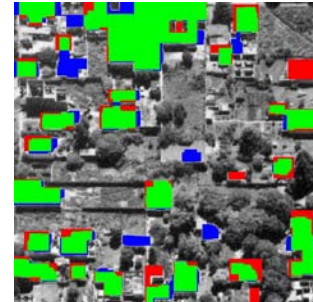


Umbral Otsu	Detectado	No Detectado	% Error
Construido	250629	5348	11.3 %
No Construido	31627	39002	





Umbral manual	Detectado	No Detectado	% Error
Construido	243526	12451	8.2%
No Construido	14494	56135	



## 5. Conclusiones

Las imágenes aéreas ofrecen mayor nitidez que las satelitales, y a pesar de no poseer colores, los resultados fueron aceptables. Por otro lado las imágenes de Google Earth son menos claras y el color no ofrece una ayuda considerable, sin embargo los resultados obtenidos en ambas fueron similares y no se justifica el gasto de adquisición de dicho servicio.

Luego de realizar las pruebas de segmentación, es posible concluir que trabajar sobre imágenes de baja calidad y sin otra información adicional, ofrece un mejor resultado el umbralizado directo que las técnicas con texturas mostradas en este trabajo, ya que es una técnica rápida y no requiere de un procesamiento riguroso. En cualquiera de los casos se debe pensar que el operador tendrá la posibilidad de mejorar la segmentación.

A futuro se pretende desarrollar un sistema que realice el segmentado mediante umbralización y permita al usuario modificar de forma manual el resultado obtenido. Además se desea que el sistema calcule las dimensiones reales escalando los resultados obtenidos. Finalmente se podrá transformar las líneas arrojadas por el proceso de segmentación en un archivo DWG. Dicho formato puede ser manipulado en aplicaciones tales como AutoCAD<sup>2</sup>.

## 6. Referencias

1. Viegas G., San Juan G., Discoli C., "Comportamiento térmico energético de Tipologías representativas pertenecientes a Mosaicos urbanos de la ciudad de La Plata." Avances en energías renovables y medio ambiente, (2007).
2. C. Lin, A. Huertas and R. Nevatia, "Detection of Buildings Using Perceptual Grouping and Shadows" (1994) Proc. IEEE Conf. Computer Vision and Pattern Recognition.
3. Chungan Lin and Ramakant Nevatiay, "Building Detection and Description from a Single Intensity Image". Computer Vision and Image Understanding: CVIU (1998)
4. Cem U" nsalan, Kim L. Boyer , "A system to detect houses and residencial street networks in multispectral satellite images. ELSEVIER Computer Vision and Image Understanding (2004)

---

<sup>2</sup> Autocad es un programa de diseño asistido por ordenador (CAD "Computer Aided Design"; en inglés, Diseño Asistido por Computador). Actualmente es desarrollado y comercializado por la empresa Autodesk. (<http://www.autodesk.es/autocad>)

5. Nilcilene das Graças Medeiros, Erivaldo Antonio da Silva, José Roberto Nogueira, “Segmentação morfológica de imagens utilizando o gradiente morfológico multi-escala”
6. D. I. Morales, M. Moctezuma, F. Parmiggiani. “Urban Edge Detection by Texture Analysis” National University of Mexico, Faculty of Engineering. (2004)
7. Aaron K. Shackelford, Student Member , “Automated 2-D Building Footprint Extraction from High-Resolution Satellite Multispectral Imagery” IEEE, Curt H. Davis, Senior Member, IEEE and Xiangyun Wang
8. G. Sohn, I. Dowman, “Building extraction using lidar data and aerial images”
9. C.R. Shyu,a,b, G. Scotta,b, M. Klarica,b, C. H. Davis,a,c, K. Palaniappan,a,b, “Automatic object extraction from full differential morphological profile in urban imagery for efficient object indexing and retrievals.
10. Marlene Rodríguez Cruz, Miguel Moctezuma Flores , “Análisis bayesiano y fusión de datos para la clasificación de escenas urbanas del distrito federal” *Ingeniería. Investigación y tecnología*. (2006).
11. Jon Atli Benediktsson, Senior Member, IEEE, Martino Pesaresi, and Kolbeinn Arnason, “Classification and Feature Extraction for Remote Sensing Images From Urban Areas Based on Morphological Transformations”
12. V. Venkateswar, Rama Chellappa, “Extraction of Straight Lines in Aerial Images” (1992) *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*,